

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Bandman, Olga
Corely, Neil C.
Shah, Purvi

(ii) TITLE OF THE INVENTION: HUMAN MITOCHONDRIAL MALATE
DEHYDROGENASE

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
(B) STREET: 3174 Porter Drive
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Herewith
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Billings, Lucy J.
(B) REGISTRATION NUMBER: 36,749
(C) REFERENCE/DOCKET NUMBER: PF-0379 US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 650-855-0555
(B) TELEFAX: 650-845-4166
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 338 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: THP1PLB01
(B) CLONE: 11587

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met	Leu	Ser	Ala	Leu	Ala	Arg	Pro	Ala	Ser	Ala	Ala	Leu	Arg	Arg	Ser
1				5					10					15	
Phe	Ser	Thr	Ser	Ala	Gln	Asn	Asn	Ala	Lys	Val	Ala	Val	Leu	Gly	Ala
			20					25					30		
Ser	Gly	Gly	Ile	Gly	Gln	Pro	Leu	Ser	Leu	Leu	Leu	Lys	Asn	Ser	Pro
		35					40					45			
Leu	Val	Ser	Arg	Leu	Thr	Leu	Tyr	Asp	Ile	Ala	His	Thr	Pro	Gly	Val
	50					55					60				
Ala	Ala	Asp	Leu	Ser	His	Ile	Glu	Thr	Lys	Ala	Ala	Val	Lys	Gly	Tyr
65					70					75					80
Leu	Gly	Pro	Glu	Gln	Leu	Pro	Asp	Cys	Leu	Lys	Gly	Cys	Asp	Val	Val
				85					90					95	
Val	Ile	Pro	Ala	Gly	Val	Pro	Arg	Lys	Pro	Gly	Met	Thr	Arg	Asp	Asp
			100					105					110		
Leu	Phe	Asn	Thr	Asn	Ala	Thr	Ile	Val	Ala	Thr	Leu	Thr	Ala	Ala	Cys
		115					120					125			
Ala	Gln	His	Cys	Pro	Glu	Ala	Met	Ile	Cys	Val	Ile	Ala	Asn	Pro	Val
	130					135					140				
Asn	Ser	Thr	Ile	Pro	Ile	Thr	Ala	Glu	Val	Phe	Lys	Lys	His	Gly	Val
145					150					155					160
Tyr	Asn	Pro	Asn	Lys	Ile	Phe	Gly	Val	Thr	Thr	Leu	Asp	Ile	Val	Arg
			165						170					175	
Ala	Asn	Thr	Phe	Val	Ala	Glu	Leu	Lys	Gly	Leu	Asp	Pro	Ala	Arg	Val
			180					185					190		
Asn	Val	Pro	Val	Ile	Gly	Gly	His	Ala	Gly	Lys	Thr	Ile	Ile	Pro	Leu
	195						200					205			
Ile	Ser	Gln	Cys	Thr	Pro	Lys	Val	Asp	Phe	Pro	Gln	Asp	Gln	Leu	Thr
	210					215					220				
Ala	Leu	Thr	Gly	Arg	Ile	Gln	Glu	Ala	Gly	Thr	Glu	Val	Val	Lys	Ala
225					230					235					240
Lys	Ala	Gly	Ala	Gly	Ser	Ala	Thr	Leu	Ser	Met	Ala	Tyr	Ala	Gly	Ala
			245						250					255	
Arg	Phe	Val	Phe	Ser	Leu	Val	Asp	Ala	Met	Asn	Gly	Lys	Glu	Gly	Val
			260					265					270		
Val	Glu	Cys	Ser	Phe	Val	Lys	Ser	Gln	Glu	Thr	Glu	Cys	Thr	Tyr	Phe
		275					280					285			
Ser	Thr	Pro	Leu	Leu	Leu	Gly	Lys	Lys	Gly	Ile	Glu	Lys	Asn	Leu	Gly
	290					295					300				
Ile	Gly	Lys	Val	Ser	Ser	Phe	Glu	Glu	Lys	Met	Ile	Ser	Asp	Ala	Ile
305					310					315					320
Pro	Glu	Leu	Lys	Ala	Ser	Ile	Lys	Lys	Gly	Glu	Asp	Phe	Val	Lys	Thr
				325					330					335	
Leu	Lys														

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1334 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: THP1PLB01
- (B) CLONE: 11587

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

GGCCCCAGAG AGCAGGCGCT GGGCAGTGTG GAGGTCGTTG GAGTCACTTC CGCGTCACCA      60
GCTCCTGTGC CTGCCAGTCG GTGCCCCCTCC CGCTCCAGCC ATGCTCTCCG CCCTCGCCCCG      120
GCCTGCCAGC GCTGCTCTCC GCCGCAGCTT CAGCACCTCG GCCCAGAACA ATGCTAAAGT      180
AGCTGTGCTA GGGGCCTCTG GAGGCATCGG GCAGCCACTT TCACTTCTCC TGAAGAACAG      240
CCCCTTGGTG AGCCGCCTGA CCTCTATGA TATCGCGCAC ACACCCGGAG TGGCCGCAGA      300
TCTGAGCCAC ATCGAGACCA AAGCCGCTGT GAAAGGCTAC CTCGGACCTG AACAGCTGCC      360
TGA CTGCTG AAAGGTTGTG ATGTGGTAGT TATTCCGGCT GGAGTCCCCA GAAAGCCAGG      420
CATGACCCGG GACGACCTGT TCAACACCAA TGCCACGATT GTGGCCACCC TGACCGCTGC      480
CTGTGCCCAG CACTGCCCCG AAGCCATGAT CTGCGTCATT GCCAATCCGG TTAATTCCAC      540
CATCCCCATC ACAGCAGAAG TT'TCAAGAA GCATGGAGTG TACAACCCCA ACAAATCTT      600
CGGCGTGACG ACCCTGGACA TCGTCAGAGC CAACACCTTT GTTGCAGAGC TGAAGGGTTT      660
GGATCCAGCT CGAGTCAACG TCCCTGTTCAT TGGTGGCCAT GCTGGGAAGA CCATCATCCC      720
CCTGATCTCT CAGTGACACC CCAAGGTGGA CTTTCCCCAG GACCAGCTGA CAGCACTCAC      780
TGGGCGGATC CAGGAGGCCG GCACGGAGGT GGTCAAGGCT AAAGCCGGAG CAGGCTCTGC      840
CACCTCTCTC ATGGCGTATG CCGGCGCCCC CTTTGTCTTC TCCCTTGTGG ATGCAATGAA      900
TGGAAAGGAA GGTGTTGTGG AATGTTCC'TT CGTTAAGTCA CAGGAAACGG AATGTACCTA      960
CTTCTCCACA CCGCTGCTGC TTGGGAAAAA GGGCATCGAG AAGAACCTGG GCATCGGCAA      1020
AGTCTCCTCT TTTGAGGAGA AGATGATCTC GGATGCCATC CCCGAGCTGA AGGCCTCCAT      1080
CAAGAAGGGG GAAGATTTTCG TGAAGACCCT GAAGTGAGCC GCTGTGACGG GTGGCCAGTT      1140
TCCTTAATTT ATGAAGGCAT CATGTCAC TG CAAAGCCGTT GCAGATAAAC TTTGTATTTT      1200
AATTTGCTTT GGTGATGATT ACTGTATTGA CATCATCATG CCTTCCAAAT TGTGGGTGGC      1260
TCTGTGGGCG CATCAATAAA AGCCGTCCTT GATTTTATTT TTCAAGGTCC CTTCTGTAAA      1320
TGCAAAAAAA AAAA                                     1334

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 338 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 56643

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Met Leu Ser Ala Leu Ala Arg Pro Val Gly Ala Ala Leu Arg Arg Ser
 1          5          10          15
Phe Ser Thr Ser Ala Gln Asn Asn Ala Lys Val Ala Val Leu Gly Ala
 20          25          30
Ser Gly Gly Ile Gly Gln Pro Leu Ser Leu Leu Leu Lys Asn Ser Pro
 35          40          45
Leu Val Ser Arg Leu Thr Leu Tyr Asp Ile Ala His Thr Pro Gly Val
 50          55          60
Ala Ala Asp Leu Ser His Ile Glu Thr Arg Ala Asn Val Lys Gly Tyr
 65          70          75          80
Leu Gly Pro Glu Gln Leu Pro Asp Cys Leu Lys Gly Cys Asp Val Val
 85          90          95
Val Ile Pro Ala Gly Val Pro Arg Lys Pro Gly Met Thr Arg Asp Asp
100          105          110
Leu Phe Asn Thr Asn Ala Thr Ile Val Ala Thr Leu Thr Ala Ala Cys
115          120          125
Ala Gln His Cys Pro Glu Ala Met Ile Cys Ile Ile Ser Asn Pro Val
130          135          140
Asn Ser Thr Ile Pro Ile Thr Ala Glu Val Phe Lys Lys His Gly Val
145          150          155          160
Tyr Asn Pro Asn Lys Ile Phe Gly Val Thr Thr Leu Asp Ile Val Arg
165          170          175

```

Ala	Asn	Thr	Phe	Val	Ala	Glu	Leu	Lys	Gly	Leu	Asp	Pro	Ala	Arg	Val
			180					185					190		
Asn	Val	Pro	Val	Ile	Gly	Gly	His	Ala	Gly	Lys	Thr	Ile	Ile	Pro	Leu
		195					200					205			
Ile	Ser	Gln	Cys	Thr	Pro	Lys	Val	Asp	Phe	Pro	Gln	Asp	Gln	Leu	Ala
	210					215					220				
Thr	Leu	Thr	Gly	Lys	Ile	Gln	Glu	Ala	Gly	Thr	Glu	Val	Val	Lys	Ala
225					230					235				240	
Lys	Ala	Gly	Ala	Gly	Ser	Ala	Thr	Leu	Ser	Met	Ala	Tyr	Ala	Gly	Ala
			245						250					255	
Arg	Phe	Val	Phe	Ser	Leu	Val	Asp	Ala	Met	Asn	Gly	Lys	Glu	Gly	Val
		260					265						270		
Ile	Glu	Cys	Ser	Phe	Val	Gln	Ser	Lys	Glu	Thr	Glu	Cys	Thr	Tyr	Phe
	275					280					285				
Ser	Thr	Pro	Leu	Leu	Leu	Gly	Lys	Lys	Gly	Leu	Glu	Lys	Asn	Leu	Gly
	290				295						300				
Ile	Gly	Lys	Ile	Thr	Pro	Phe	Glu	Glu	Lys	Met	Ile	Ala	Glu	Ala	Ile
305					310					315					320
Pro	Glu	Leu	Lys	Ala	Ser	Ile	Lys	Lys	Gly	Glu	Asp	Phe	Val	Lys	Asn
				325					330					335	

Met Lys

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 298 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 164541

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ser	Leu	Leu	Leu	Lys	Asn	Ser	Pro	Leu	Val	Ser	Arg	Leu	Thr	Leu	Tyr
1				5					10					15	
Asp	Ile	Ala	His	Thr	Pro	Gly	Val	Ala	Ala	Asp	Leu	Ser	His	Ile	Glu
		20						25					30		
Thr	Arg	Ala	Thr	Val	Lys	Gly	Tyr	Leu	Gly	Pro	Glu	Gln	Leu	Pro	Asp
		35				40						45			
Cys	Leu	Lys	Gly	Cys	Asp	Val	Val	Val	Ile	Pro	Ala	Gly	Val	Pro	Arg
	50				55					60					
Lys	Pro	Gly	Met	Thr	Arg	Asp	Asp	Leu	Phe	Asn	Thr	Asn	Ala	Thr	Met
65					70					75				80	
Val	Ala	Thr	Leu	Thr	Val	Ala	Cys	Ala	Gln	His	Cys	Pro	Asp	Ala	Met
			85						90					95	
Ile	Cys	Ile	Ile	Ser	Asn	Pro	Val	Asn	Ser	Thr	Ile	Pro	Met	Thr	Ala
		100						105					110		
Glu	Val	Phe	Lys	Lys	His	Gly	Val	Tyr	Asn	Pro	Asn	Lys	Ile	Phe	Gly
		115				120						125			
Val	Thr	Thr	Leu	Asp	Ile	Val	Arg	Ala	Asn	Ala	Phe	Val	Ala	Glu	Leu
		130				135					140				
Lys	Gly	Leu	Asp	Pro	Ala	Arg	Val	Ser	Val	Pro	Val	Ile	Gly	Gly	His
145					150					155				160	
Ala	Gly	Lys	Thr	Ile	Ile	Pro	Leu	Ile	Ser	Gln	Cys	Thr	Pro	Lys	Val
				165					170					175	

Asp	Phe	Pro	Gln	Asp	Gln	Leu	Ser	Thr	Leu	Thr	Gly	Arg	Ile	Gln	Glu
			180					185					190		
Ala	Gly	Thr	Glu	Val	Val	Lys	Ala	Lys	Ala	Gly	Ala	Gly	Ser	Ala	Thr
		195					200					205			
Leu	Ser	Met	Ala	Tyr	Ala	Gly	Ala	Arg	Phe	Val	Phe	Ser	Leu	Val	Asp
	210					215					220				
Ala	Met	Asn	Gly	Lys	Glu	Gly	Val	Val	Glu	Cys	Ser	Phe	Val	Lys	Ser
225					230					235				240	
Gln	Glu	Thr	Asp	Cys	Pro	Tyr	Phe	Ser	Thr	Pro	Leu	Leu	Leu	Gly	Lys
			245						250					255	
Lys	Gly	Ile	Glu	Lys	Asn	Leu	Arg	Ile	Gly	Lys	Ile	Ser	Pro	Phe	Glu
		260						265					270		
Glu	Lys	Met	Ile	Ala	Glu	Ala	Ile	Pro	Glu	Leu	Lys	Ala	Ser	Ile	Lys
		275					280					285			
Lys	Gly	Glu	Glu	Phe	Val	Lys	Asn	Thr	Lys						
	290					295									